

## Raw Sequence Listing Error Summary

**ERROR DETECTED**

### SUGGESTED CORRECTION

SERIAL NUMBER: 04/889, 767

**ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE**

- 1 \_\_\_\_\_ Wrapped Nucleics  
Wrapped Aminos: The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 \_\_\_\_\_ Invalid Line Length The rules require that a line **not exceed** 72 characters in length. This includes white spaces.
- 3 \_\_\_\_\_ Misaligned Amino  
Numbering The numbering under each 5<sup>th</sup> amino acid is misaligned. Do **not** use tab codes between numbers; use **space characters**, instead.
- 4 \_\_\_\_\_ Non-ASCII The submitted file was **not** saved in ASCII(DOS) text, as **required** by the Sequence Rules. Please **ensure your subsequent submission is saved in ASCII text**.
- 5 \_\_\_\_\_ Variable Length Sequence(s) \_\_\_\_\_ contain n's or Xaa's representing more than one residue. **Per Sequence Rules, each n or Xaa can only represent a single residue.** Please present the **maximum** number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 \_\_\_\_\_ PatentIn 2.0  
"bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) \_\_\_\_\_. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**
- 7 \_\_\_\_\_ Skipped Sequences  
(OLD RULES) Sequence(s) \_\_\_\_\_ missing. If intentional, please insert the following lines for **each** skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped
- Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences.
- 8 \_\_\_\_\_ Skipped Sequences  
(NEW RULES) Sequence(s) \_\_\_\_\_ missing. If intentional, please insert the following lines for **each** skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 9 \_\_\_\_\_ Use of n's or Xaa's  
(NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is **MANDATORY** if n's or Xaa's are present.  
In <220> to <223> section, please explain location of **n** or **Xaa**, and which residue **n** or **Xaa** represents.
- 10 \_\_\_\_\_ Invalid <213>  
Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is **required** when <213> response is Unknown or is Artificial Sequence
- 11 \_\_\_\_\_ Use of <220> Sequence(s) \_\_\_\_\_ missing the <220> "Feature" and associated numeric identifiers and responses.  
Use of <220> to <223> is **MANDATORY** if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 \_\_\_\_\_ PatentIn 2.0  
"bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

O1PE

## RAW SEQUENCE LISTING

DATE: 07/24/2001

PATENT APPLICATION: US/09/884,767

TIME: 07:52:05

Input Set : A:\DYX-012.1 US seqlist.txt

Output Set: N:\CRF3\07242001\I884767.raw

3 <110> APPLICANT: DYAX Corp.  
 4 Ley, Arthur C.  
 5 Luneau, Christopher J.  
 6 Ladner, Robert C  
 8 <120> TITLE OF INVENTION: NOVEL ENTEROKINASE CLEAVAGE SEQUENCES  
 10 <130> FILE REFERENCE: DYX-012.1 US, DYX-012.1 PCT  
 C--> 12 <140> CURRENT APPLICATION NUMBER: US/09/884,767  
 13 <141> CURRENT FILING DATE: 2001-06-19  
 15 <150> PRIOR APPLICATION NUMBER: US 09/597,321  
 16 <151> PRIOR FILING DATE: 2000-06-19  
 18 <160> NUMBER OF SEQ ID NOS: 217  
 20 <170> SOFTWARE: PatentIn version 3.1  
 22 <210> SEQ ID NO: 1  
 23 <211> LENGTH: 9  
 24 <212> TYPE: PRT  
 25 <213> ORGANISM: synthetic enterokinase cleavage sequence  
 27 <220> FEATURE:  
 28 <221> NAME/KEY: MISC\_FEATURE  
 29 <222> LOCATION: (1)..(1)  
 30 <223> OTHER INFORMATION: Xaa1 is an optional polypeptide of one or more amino acids  
 33 <220> FEATURE:  
 34 <221> NAME/KEY: MISC\_FEATURE  
 35 <222> LOCATION: (2)..(2)  
 36 <223> OTHER INFORMATION: Xaa2 is an optional amino acid which, if present, is Ala,  
 Asp, Glu,  
 37 u, Phe, Gly, Ile, Asn, Ser, or Val  
 40 <220> FEATURE:  
 41 <221> NAME/KEY: MISC\_FEATURE  
 42 <222> LOCATION: (3)..(3)  
 43 <223> OTHER INFORMATION: Xaa3 is an optional amino acid which, if present, is Ala,  
 Asp, Glu,  
 44 u, His, Ile, Leu, Met, Gln or Ser  
 47 <220> FEATURE:  
 48 <221> NAME/KEY: MISC\_FEATURE  
 49 <222> LOCATION: (4)..(4)  
 50 <223> OTHER INFORMATION: Xaa4 is an optional amino acid which, if present, is Asp,  
 Glu, Ph  
 51 e, His, Ile, Met, Asn, Pro, Val, or Trp  
 54 <220> FEATURE:  
 55 <221> NAME/KEY: MISC\_FEATURE  
 56 <222> LOCATION: (5)..(5)  
 57 <223> OTHER INFORMATION: Xaa5 is Ala, Asp, Glu, or Thr  
 60 <220> FEATURE:  
 61 <221> NAME/KEY: MISC\_FEATURE  
 62 <222> LOCATION: (8)..(8)  
 63 <223> OTHER INFORMATION: Xaa8 is any amino acid  
 66 <220> FEATURE:  
 67 <221> NAME/KEY: MISC\_FEATURE  
 68 <222> LOCATION: (9)..(9)  
 69 <223> OTHER INFORMATION: Xaa9 is an optional polypeptide of at least one amino acid

Does Not Comply  
 Corrected Diskette Needed  
 See page 2 & 5

See item #10  
 on ERROR Summary  
 SHEET.

Xaa can only be a single amino acid

each Xaa can only be equal to one amino acid.

## RAW SEQUENCE LISTING

DATE: 07/24/2001

PATENT APPLICATION: US/09/884,767

TIME: 07:52:05

Input Set : A:\DYX-012.1 US seqlist.txt

Output Set: N:\CRF3\07242001\I884767.raw

72 <400> SEQUENCE: 1  
W--> 74 Xaa Xaa Xaa Xaa Xaa Asp Arg Xaa Xaa  
75 1 5  
78 <210> SEQ ID NO: 2  
79 <211> LENGTH: 9  
80 <212> TYPE: PRT  
81 <213> ORGANISM: synthetic enterokinase cleavage sequence  
83 <220> FEATURE:  
84 <221> NAME/KEY: MISC\_FEATURE  
85 <222> LOCATION: (1)..(1)  
86 <223> OTHER INFORMATION: Xaa1 is an optional polypeptide of one or more amino acids  
89 <220> FEATURE:  
90 <221> NAME/KEY: MISC\_FEATURE  
91 <222> LOCATION: (2)..(2)  
92 <223> OTHER INFORMATION: Xaa2 is an optional amino acid which, if present, is Asp or  
Glu  
95 <220> FEATURE:  
96 <221> NAME/KEY: MISC\_FEATURE  
97 <222> LOCATION: (3)..(3)  
98 <223> OTHER INFORMATION: Xaa3 is an optional amino acid which, if present, is Val  
101 <220> FEATURE:  
102 <221> NAME/KEY: MISC\_FEATURE  
103 <222> LOCATION: (4)..(4)  
104 <223> OTHER INFORMATION: Xaa4 is an optional amino acid which, if present, is Tyr  
107 <220> FEATURE:  
108 <221> NAME/KEY: MISC\_FEATURE  
109 <222> LOCATION: (5)..(5)  
110 <223> OTHER INFORMATION: Xaa5 is Asp, Glu or Ser  
113 <220> FEATURE:  
114 <221> NAME/KEY: MISC\_FEATURE  
115 <222> LOCATION: (8)..(8)  
116 <223> OTHER INFORMATION: Xaa8 is any amino acid  
119 <220> FEATURE:  
120 <221> NAME/KEY: MISC\_FEATURE  
121 <222> LOCATION: (9)..(9)  
122 <223> OTHER INFORMATION: Xaa9 is an optional polypeptide of one or more amino acid  
125 <400> SEQUENCE: 2  
W--> 127 Xaa Xaa Xaa Xaa Xaa Glu Arg Xaa Xaa  
128 1 5  
131 <210> SEQ ID NO: 3  
132 <211> LENGTH: 7  
133 <212> TYPE: PRT  
134 <213> ORGANISM: synthetic enterokinase cleavage sequence  
136 <220> FEATURE:  
137 <221> NAME/KEY: MISC\_FEATURE  
138 <222> LOCATION: (7)..(7)  
139 <223> OTHER INFORMATION: Xaa is any amino acid  
142 <400> SEQUENCE: 3  
W--> 144 Asp Ile Asn Asp Asp Arg Xaa  
145 1 5

The types of errors shown exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/884,767

DATE: 07/24/2001

TIME: 07:52:05

Input Set : A:\DYX-012.1 US seqlist.txt

Output Set: N:\CRF3\07242001\I884767.raw

148 <210> SEQ ID NO: 4  
 149 <211> LENGTH: 7  
 150 <212> TYPE: PRT  
 151 <213> ORGANISM: synthetic enterokinase cleavage sequence  
 153 <220> FEATURE:  
 154 <221> NAME/KEY: MISC\_FEATURE  
 155 <222> LOCATION: (7)..(7)  
 156 <223> OTHER INFORMATION: Xaa is any amino acid  
 159 <400> SEQUENCE: 4  
 W--> 161 Gly Asn Tyr Thr Asp Arg Xaa  
 162 1 5  
 165 <210> SEQ ID NO: 5  
 166 <211> LENGTH: 6  
 167 <212> TYPE: PRT  
 168 <213> ORGANISM: streptavidin binding sequence — see page 1  
 170 <400> SEQUENCE: 5  
 172 Cys His Pro Gln Phe Cys  
 173 1 5  
 176 <210> SEQ ID NO: 6  
 177 <211> LENGTH: 4  
 178 <212> TYPE: PRT  
 179 <213> ORGANISM: streptavidin binding sequence  
 181 <400> SEQUENCE: 6  
 183 His Pro Gln Phe  
 184 1  
 187 <210> SEQ ID NO: 7  
 188 <211> LENGTH: 9  
 189 <212> TYPE: PRT  
 190 <213> ORGANISM: streptavidin binding sequence  
 192 <400> SEQUENCE: 7  
 194 Cys His Pro Gln Phe Cys Ser Trp Arg  
 195 1 5  
 198 <210> SEQ ID NO: 8  
 199 <211> LENGTH: 6  
 200 <212> TYPE: PRT  
 201 <213> ORGANISM: enterokinase cleavage sequence  
 203 <220> FEATURE:  
 204 <221> NAME/KEY: MISC\_FEATURE  
 205 <222> LOCATION: (6)..(6)  
 206 <223> OTHER INFORMATION: Xaa is Ile (natural trypsinogen site) or any amino acid  
 (syntheti  
 207 c cleavage sites)  
 210 <400> SEQUENCE: 8  
 W--> 212 Asp Asp Asp Asp Lys Xaa  
 213 1 5  
 216 <210> SEQ ID NO: 9  
 217 <211> LENGTH: 86  
 218 <212> TYPE: PRT  
 219 <213> ORGANISM: exogenous display polypeptide of a phage display library  
 221 <220> FEATURE:

## RAW SEQUENCE LISTING

DATE: 07/24/2001

PATENT APPLICATION: US/09/884,767

TIME: 07:52:05

Input Set : A:\DYX-012.1 US seqlist.txt

Output Set: N:\CRF3\07242001\I884767.raw

222 <221> NAME/KEY: MISC\_FEATURE  
 223 <222> LOCATION: (43)..(55)  
 224 <223> OTHER INFORMATION: X is any amino acid except Cys  
 227 <400> SEQUENCE: 9  
 229 Ala Glu Trp His Pro Gln Phe Ser Ser Pro Ser Ala Ser Arg Pro Ser  
 230 1 5 10 15  
 233 Glu Gly Pro Cys His Pro Gln Phe Pro Arg Cys Tyr Ile Glu Asn Leu  
 234 20 25 30  
 W--> 237 Asp Glu Phe Arg Pro Gly Gly Ser Gly Gly Xaa Xaa Xaa Xaa Xaa Xaa  
 238 35 40 45  
 W--> 241 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Ala Gln Ser Asp Gly Gly Gly Ser  
 242 50 55 60  
 245 Thr Glu His Ala Glu Gly Gly Ser Ala Asp Pro Ser Tyr Ile Glu Gly  
 246 65 70 75 80  
 249 Arg Ile Val Gly Ser Ala  
 250 85  
 253 <210> SEQ ID NO: 10  
 254 <211> LENGTH: 7  
 255 <212> TYPE: PRT  
 256 <213> ORGANISM: synthetic enterokinase cleavage sequence  
 258 <400> SEQUENCE: 10  
 260 Tyr Glu Trp Gln Asp Arg Thr  
 261 1 5  
 264 <210> SEQ ID NO: 11  
 265 <211> LENGTH: 7  
 266 <212> TYPE: PRT  
 267 <213> ORGANISM: synthetic enterokinase cleavage sequence  
 269 <400> SEQUENCE: 11  
 271 Asn Ser Ile Lys Asp Arg Val  
 272 1 5  
 275 <210> SEQ ID NO: 12  
 276 <211> LENGTH: 7  
 277 <212> TYPE: PRT  
 278 <213> ORGANISM: synthetic enterokinase cleavage sequence  
 280 <400> SEQUENCE: 12  
 282 Ala Lys Ala Thr Glu Arg His  
 283 1 5  
 286 <210> SEQ ID NO: 13  
 287 <211> LENGTH: 7  
 288 <212> TYPE: PRT  
 289 <213> ORGANISM: synthetic enterokinase cleavage sequence  
 291 <400> SEQUENCE: 13  
 293 Leu Gly Lys Val Asp Arg Thr  
 294 1 5  
 297 <210> SEQ ID NO: 14  
 298 <211> LENGTH: 7  
 299 <212> TYPE: PRT  
 300 <213> ORGANISM: synthetic enterokinase cleavage sequence  
 302 <400> SEQUENCE: 14

See page 1

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/884,767

DATE: 07/24/2001

TIME: 07:52:05

Input Set : A:\DYX-012.1 US seqlist.txt

Output Set: N:\CRF3\07242001\I884767.raw

```

304 Gly Gly Met Ala Asp Lys Phe
305 1 5
308 <210> SEQ ID NO: 15
309 <211> LENGTH: 7
310 <212> TYPE: PRT
311 <213> ORGANISM: synthetic enterokinase cleavage sequence
313 <400> SEQUENCE: 15
315 Gly His Trp Leu Asp Lys Asn
316 1 5
319 <210> SEQ ID NO: 16
320 <211> LENGTH: 7
321 <212> TYPE: PRT
322 <213> ORGANISM: synthetic enterokinase cleavage sequence
324 <400> SEQUENCE: 16
326 Asn Lys Ala Lys Asp Arg Met
327 1 5
330 <210> SEQ ID NO: 17
331 <211> LENGTH: 7
332 <212> TYPE: PRT
333 <213> ORGANISM: synthetic enterokinase cleavage sequence
335 <400> SEQUENCE: 17
337 Ser Glu Asn Phe Asp Lys Asn
338 1 5
341 <210> SEQ ID NO: 18
342 <211> LENGTH: 7
343 <212> TYPE: PRT
344 <213> ORGANISM: synthetic enterokinase cleavage sequence
346 <400> SEQUENCE: 18
348 Leu Asp Trp Glu Asp Arg Ala
349 1 5
352 <210> SEQ ID NO: 19
353 <211> LENGTH: 7
354 <212> TYPE: PRT
355 <213> ORGANISM: synthetic enterokinase cleavage sequence
357 <400> SEQUENCE: 19
359 Ser Thr Asp Ala Glu Arg Met
360 1 5
363 <210> SEQ ID NO: 20
364 <211> LENGTH: 7
365 <212> TYPE: PRT
366 <213> ORGANISM: synthetic enterokinase cleavage sequence
368 <400> SEQUENCE: 20
370 His Thr Phe Ser Asp Arg Gln
371 1 5
374 <210> SEQ ID NO: 21
375 <211> LENGTH: 7
376 <212> TYPE: PRT
377 <213> ORGANISM: synthetic enterokinase cleavage sequence
379 <400> SEQUENCE: 21

```

- see page 1

FYI.

**Please Note:**

Use f n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields f each sequence which presents at least one n or Xaa.

## VERIFICATION SUMMARY

DATE: 07/24/2001

PATENT APPLICATION: US/09/884,767

TIME: 07:52:06

Input Set : A:\DYX-012.1 US seqlist.txt

Output Set: N:\CRF3\07242001\I884767.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:74 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:127 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:144 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:161 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
L:212 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8  
L:237 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9  
L:241 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9  
L:2400 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:204  
L:2417 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:205  
L:2461 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:206  
L:2502 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:207